AMENDMENTS TO SPECIFICATION:

Please replace the paragraph beginning at page 18, line 28, with the following amended paragraph:

Note that the present embodiment is not limited to the specific examples described above, but various modifications are possible without departing from the spirit of the embodiment. For example, although in the above example, feature elements to be used to obtain a categorization result such as that shown in Fig. 5 are extracted from the titles of the respective documents, and clustering is performed on the basis of the feature elements extracted from the titles, what is essential to the present embodiment is that after categorizing documents into clusters according to semantic similarity, similar clusters are merged. Therefore, the manner of clustering a plurality of documents is not limited to a particular method. For example, instead of clustering documents on the basis of feature elements extracted from the titles of the documents according to the above embodiment, clustering may also be performed according to URL addresses (after removing "http://", the remaining part is employed), updated date/time (without any restriction or within last one ~~moth~~month), or file sizes (the sizes of the Web pages in bytes). One of these items or some combination of these items may be employed in the clustering process. A desired item can be selected, for example, from a menu. In the case where a selected item is not included in a document, another item may be employed instead of the selected item. For example, when the title is selected as the item, if a Web page does not include a title, a URL address may be employed.

Please replace the paragraph beginning at page 30, line 29, with the following amended paragraph:

As described above, when a cluster correlation score K calculated by the cluster merging unit 92 is equal to or greater than TH1 ($K \geq$ TH1), original clusters merged together into a final cluster can be regarded as having very close relation and being very similar to each other. Thus, in such a case, the name of the final cluster is given by a combination of the original cluster names, which are displayed by successively listing the original cluster names. In the case of "paper cluster" and "cassette cluster", the cluster names "paper" and "cassette" can be combined into "paper cassette".

Please replace the paragraph beginning at page 32, line 16, with the following amended paragraph:

In the specific example described above, the original cluster names "product", "usage", and "outline" are represented not in a simple successive fashion but in a fashion in which a delimiter is placed between adjacent cluster names such as ~~""~~ as "product·usage·outline". If a user is informed in advance that a delimiter placed between adjacent cluster names represents "OR", the user can understand that the final cluster obtained via the cluster merging process includes some documents having contents about "product", "usage", or "outline". When a final cluster name is represented in the OR form, the delimiter placed between original cluster names is not limited to a dot as is used in "produce·usage·outline" but other types of delimiters may be used. For example, "/" may be placed between original cluster names such as "product/usage/outline".